

V. Ryan

1641

PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/189,415

DATE: 03/10/2000
TIME: 15:42:36

Input Set: I189415.RAW

This Raw Listing contains the General Information
Section and up to first 5 pages.

Does Not Comply
Corrected Diskette Needed

p. 2

```
1 <110> APPLICANT: Finlay, Brett B
2      Kenny, Brendan
3      DeVinney, Rebekah
4      Stein, Markus
5 <120> TITLE OF INVENTION: HOST RECEPTOR FOR PATHOGENIC BACTERIA
6 <130> FILE REFERENCE: 07422/013001
7 <140> CURRENT APPLICATION NUMBER: US/09/189,415
8 <141> CURRENT FILING DATE: 1998-11-10
9 <150> EARLIER APPLICATION NUMBER: 60/065,130
10 <151> EARLIER FILING DATE: 1997-11-12
11 <160> NUMBER OF SEQ ID NOS: 9
12 <170> SOFTWARE: PatentIn Ver. 2.0
13 <210> SEQ ID NO 1
14 <211> LENGTH: 1920
15 <212> TYPE: DNA
16 <213> ORGANISM: Escherichia coli
17 <400> SEQUENCE: 1
18      cggctgcata ccattacgtc atagtaatat aaaggaacgt gtcaaatttc taaataaaag 60
19      gatatatgta tgcctattgg taaccttgggt aataatgtaa atggcaatca tttaattccc 120
20      cctgcgccgc cactaccttc acaaacagac ggcgcggcac ggggaggaac tgggtcatcta 180
21      attagctcta caggagcatt aggatctcgt tcattgtttt ctcccctgag aaattctatg 240
22      gctgattctg tcgattccag agatattcca ggacttccta caaacccatc gaggcttgct 300
23      gcagctacat ctgagacatg cttgcttgga ggatttgaag ttctccatga taaggggcca 360
24      cttgatattc tcaatacgca aattggaccc tctgcatttc gtgttgaagt gcaggcagat 420
25      ggtactcatg ccgctattgg agaaaaaaat ggtttgaggg ttagcggtac attaagtcct 480
26      caagaatgga gcagcttgca atctattgat actgagggtg aaaacagatt tgtttttacc 540
27      gggggacgtg gcggtagtgg gcatccgatg gtcactgtcg catcagatat cgcggaagct 600
28      cgtacgaaaa tactggccaa attagaccca gacaatcatg gaggacgtca acccaaggac 660
29      gttgatacgc gttctgttgg tgttggcagc gcttcgggaa tagatgatgg cgttgtagc 720
30      gaaaccata cttcaacaac aaattccagc gttcgctcag atcctaaatt ctgggtttct 780
31      gtcggcgcaa ttgctgctgg tttagcggga ctggcgcaa ctggtattgc acaggcgttg 840
32      gctttgacac cggaaccgga tgatcctaca accaccgatc ctgatcaggc cgcaaagtct 900
33      gcagaaaagt caacaaaaga tcagttaacg caagaagcat tcaagaaccc tgagaaccag 960
34      aaagttaaca tcgatgcaaa cggaatgct attccgtctg gggaattaaa agatgatatt 1020
35      gttgagcaaa tagcacaaca agctaaagag gctggtgagg tggccagaca gcaggctgtt 1080
36      gaaagcaatg cacaggcgca gcagcgatat gaggatcagc atgccagacg tcaggaggaa 1140
37      ttacagcttt catcgggtat tggttacggc ctcagcagtg cattgattgt tgctggggga 1200
38      attggtgctg gtgtaacgac tgcgtccat agacgaaatc agccggcaga acagacaact 1260
39      actacaacaa cacatacggg agtgcagcaa cagaccggag ggatacccca gcacaagggtg 1320
40      gcaactgatg cacaagagcg aagacgcttc tctgatagac gtgattcgca ggggagtgtt 1380
41      gcatcgacac actggtcaga ttctctagc gaagtgggta atccatatgc tgaagttggg 1440
42      ggggctcgga atagtctatc ggctcatcag ccagaagagc atatttatga tgaggtcgct 1500
43      gcagatcctg gttatagcgt tattcagaat ttttcaggga gcggcccagt taccggaagg 1560
44      ttaataggaa ctccagggca aggtatccaa agtacttatg cgcttctggc aaacagcggc 1620
```

PAGE: 2

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/189,415

DATE: 03/10/2000
TIME: 15:42:36

Input Set: I189415.RAW

```

45      ggattgcgtt taggtatggg aggattaacg agtgggtggcg agacggcagt aagttctgta 1680
46      aatgccgcac caacgcaggg accagtacgt ttcgtttaaa tatactctgtg agtatttagt 1740
47      tgaggttggg gtgggggtggg ggggcgtttt actagcgtta atgtttcaga gaacaacgtt 1800
48      gcagcatggg taactcttga acttctgtta ttataatcaa ttaagagaaa ttataatgtc 1860
49      atcaagatat gaacttttat tagataggtt tgcggaaaaa attggtgttg gatctatttc 1920
50      <210> SEQ ID NO 2
51      <211> LENGTH: 549
52      <212> TYPE: PRT
53      <213> ORGANISM: Escherichia coli
54      <400> SEQUENCE: 2
55      Met Pro Ile Gly Asn Leu Gly Asn Asn Val Asn Gly Asn His Leu Ile
56      1          5          10          15
57      Pro Pro Ala Pro Pro Leu Pro Ser Gln Thr Asp Gly Ala Ala Arg Gly
58      20          25          30
59      Gly Thr Gly His Leu Ile Ser Ser Thr Gly Ala Leu Gly Ser Arg Ser
60      35          40          45
61      Leu Phe Ser Pro Leu Arg Asn Ser Met Ala Asp Ser Val Asp Ser Arg
62      50          55          60
63      Asp Ile Pro Gly Leu Pro Thr Asn Pro Ser Arg Leu Ala Ala Ala Thr
64      65          70          75          80
65      Ser Glu Thr Cys Leu Leu Gly Gly Phe Glu Val Leu His Asp Lys Gly
66      85          90          95
67      Pro Leu Asp Ile Leu Asn Thr Gln Ile Gly Pro Ser Ala Phe Arg Val
68      100         105         110
69      Glu Val Gln Ala Asp Gly Thr His Ala Ala Ile Gly Glu Lys Asn Gly
70      115         120         125
71      Leu Glu Val Ser Val Thr Leu Ser Pro Gln Glu Trp Ser Ser Leu Gln
72      130         135         140
73      Ser Ile Asp Thr Glu Gly Lys Asn Arg Phe Val Phe Thr Gly Gly Arg
74      145         150         155         160
75      Gly Gly Ser Gly His Pro Met Val Thr Val Ala Ser Asp Ile Ala Glu
76      165         170         175
77      Ala Arg Thr Arg Ile Leu Ala Lys Leu Asp Pro Asp Asn His Gly Gly
78      180         185         190
79      Arg Gln Pro Lys Asp Val Asp Thr Arg Ser Val Gly Val Gly Ser Ala
80      195         200         205
81      Ser Gly Ile Asp Asp Gly Val Val Ser Glu Thr His Thr Ser Thr Thr
82      210         215         220
83      Asn Ser Ser Val Arg Ser Asp Pro Lys Phe Trp Val Ser Val Gly Ala
84      225         230         235         240
85      Ile Ala Ala Gly Leu Ala Gly Leu Ala Ala Thr Gly Ile Ala Gln Ala
86      245         250         255
87      Leu Ala Leu Thr Pro Glu Pro Asp Asp Pro Thr Thr Thr Asp Pro Asp
88      260         265         270
89      Gln Ala Ala Asn Ala Ala Glu Ser Ala Thr Lys Asp Gln Leu Thr Gln
90      275         280         285
91      Glu Ala Phe Lys Asn Pro Glu Asn Gln Lys Val Asn Ile Asp Ala Asn
92      290         295         300
93      Gly Asn Ala Ile Pro Ser Gly Glu Leu Xaa Asp Asp Ile Val Glu Gln
94      305         310         315         320

```

W-->

See item 10
on Ena summary
sheet

PAGE: 3

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/189,415

DATE: 03/10/2000
TIME: 15:42:36

Input Set: I189415.RAW

```

95   Ile Ala Gln Gln Ala Lys Glu Ala Gly Glu Val Ala Arg Gln Gln Ala
96                               325                               330           335
97   Val Glu Ser Asn Ala Gln Ala Gln Gln Arg Tyr Glu Asp Gln His Ala
98                               340                               345           350
99   Arg Arg Gln Glu Glu Leu Gln Leu Ser Ser Gly Ile Gly Tyr Gly Leu
100                              355                               360           365
101   Ser Ser Ala Leu Ile Val Ala Gly Gly Ile Gly Ala Gly Val Thr Thr
102                              370                               375           380
103   Ala Leu His Arg Arg Asn Gln Pro Ala Glu Gln Thr Thr Thr Thr Thr
104                              385                               390           395           400
105   Thr His Thr Val Val Gln Gln Gln Thr Gly Gly Ile Pro Gln His Lys
106                              405                               410           415
107   Val Ala Leu Met Pro Gln Glu Arg Arg Arg Phe Ser Asp Arg Arg Asp
108                              420                               425           430
109   Ser Gln Gly Ser Val Ala Ser Thr His Trp Ser Asp Ser Ser Ser Glu
110                              435                               440           445
111   Val Val Asn Pro Tyr Ala Glu Val Gly Gly Ala Arg Asn Ser Leu Ser
112                              450                               455           460
113   Ala His Gln Pro Glu Glu His Ile Tyr Asp Glu Val Ala Ala Asp Pro
114                              465                               470           475           480
115   Gly Tyr Ser Val Ile Gln Asn Phe Ser Gly Ser Gly Pro Val Thr Gly
116                              485                               490           495
117   Arg Leu Ile Gly Thr Pro Gly Gln Gly Ile Gln Ser Thr Tyr Ala Leu
118                              500                               505           510
119   Leu Ala Asn Ser Gly Gly Leu Arg Leu Gly Met Gly Gly Leu Thr Ser
120                              515                               520           525
121   Gly Gly Glu Thr Ala Val Ser Ser Val Asn Ala Ala Pro Thr Pro Gly
122                              530                               535           540
123   Pro Val Arg Phe Val
124                              545

```

<210> SEQ ID NO 3

<211> LENGTH: 1723

<212> TYPE: DNA

<213> ORGANISM: Escherichia coli

<400> SEQUENCE: 3

```

130   atgcctattg gtaaccttgg tcataatccc aatgtgaata attcaattcc tcctgcacct 60
131   ccattacctt cacaaaccga cgggtgcagg gggcgtgggtc agctcattaa ctctacgggg 120
132   ccgttgggat ctcgtgcgct atttacgcct gtaaggaatt ctatggctga ttctggcgac 180
133   aatcgtgcca gtgatgttcc tggacttcct gtaaattccga tgcgcctggc ggcgtctgag 240
134   ataacactga atgatggatt tgaagttcct catgatcatg gtccgctcga tactcttaac 300
135   aggagattg gctcttcggt atttcagatt gaaactcagg aagatggtaa acatattgct 360
136   gtcggtcaga ggaatggtgt tgagacctct gttgttttaa gtgatcaaga gtacgctcgc 420
137   ttgcagtcca ttgatcctga aggtaaagac aaatttgtat ttactggagg ccgtggtggt 480
138   gctgggcatg ctatggtcac cgttgcttca gatatcacgg aagcccgcc aaggatactg 540
139   gagctgttag agcccaaagg gaccggggag tccaaagggt ctggggagtc aaaaggcgtt 600
140   ggggagttag gggagtcaaa tagcgggtgc gaaaacacca cagaaactca gacctcaacc 660
141   tcaacttcca gccttcgttc agatcctaaa ctttggttgg cgttggggac tgttgctaca 720
142   ggtctgatag ggttggcggc gacgggtatt gtacaggcgc ttgcattgac gccggagccg 780
143   gatagcccaa ccacgaccga ccctgatgca gctgcaagtg caactgaaac tgcgacaaga 840
144   gatcagttaa cgaaagaagc gttccagaac ccagataatc aaaaagttaa tatcgatgag 900

```

PAGE: 4

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/189,415

DATE: 03/10/2000
TIME: 15:42:36

Input Set: I189415.RAW

```

145      ctccgaaatg cgattccgtc aggggtattg aaagatgatg ttgttgcgaa tatagaagag 960
146      caggctaaaag cagcaggcga agaggccaaa cagcaagcca ttgaaaataa tgctcaggcg 1020
147      caaaaaaaaa atgatgaaca acaagctaaa cgccaggagg agctgaaagt ttcacgggg 1080
148      gctggctacg gtcttagtgg cgcattgatt cttgggtggg gaattggtgt tgccgtcacc 1140
149      gctgcgcttc atcgaaaaaa tcagccggta gaacaaacaa caacaactac tactacaact 1200
150      acaactacaa gcgcacgtac ggtagagaat aagcctgcaa ataatacacc tgcacagggc 1260
151      aatgtagata cccctgggtc agaagatacc atggagagca gacgtagctc gatggctagc 1320
152      acctcgctga ctttctttga cacttccagc atagggaccg tgcagaatcc gtatgctgat 1380
153      gttaaaacat cgctgcatga ttcgcagggt ccgacttcta attctaatac gtctgttcag 1440
154      aatatgggga atacagattc tgttgatat agcaccattc aacatcctcc ccgggatact 1500
155      actgataacg gcgcacgggt attaggaaat ccaagtgcgg ggattcaaag cacttatgcg 1560
156      cgtctggcgc taagtgggtg attacgccat gacatgggag gattaacggg ggggagtaat 1620
157      agcgtgtgta atacttcgaa taaccaccca gcgccgggat cccatcgttt cgtctaaata 1680
158      tatccataat cattttattt agaggggagg agggggggaag tct 1723

```

159 <210> SEQ ID NO 4

160 <211> LENGTH: 559

161 <212> TYPE: PRT

162 <213> ORGANISM: Escherichia coli

163 <400> SEQUENCE: 4

```

164      Met Pro Ile Gly Asn Leu Gly His Asn Pro Asn Val Asn Asn Ser Ile
165      1          5          10          15
166      Pro Pro Ala Pro Pro Leu Pro Ser Gln Thr Asp Gly Ala Gly Gly Arg
167      20          25          30
168      Gly Gln Leu Ile Asn Ser Thr Gly Pro Leu Gly Ser Arg Ala Leu Phe
169      35          40          45
170      Thr Pro Val Arg Asn Ser Met Ala Asp Ser Gly Asp Asn Arg Ala Ser
171      50          55          60
172      Asp Val Pro Gly Leu Pro Val Asn Pro Met Arg Leu Ala Ala Ser Glu
173      65          70          75          80
174      Ile Thr Leu Asn Asp Gly Phe Glu Val Leu His Asp His Gly Pro Leu
175      85          90          95
176      Asp Thr Leu Asn Arg Gln Ile Gly Ser Ser Val Phe Arg Val Glu Thr
177      100         105         110
178      Gln Glu Asp Gly Lys His Ile Ala Val Gly Gln Arg Asn Gly Val Glu
179      115         120         125
180      Thr Ser Val Val Leu Ser Asp Gln Glu Tyr Ala Arg Leu Gln Ser Ile
181      130         135         140
182      Asp Pro Glu Gly Lys Asp Lys Phe Val Phe Thr Gly Gly Arg Gly Gly
183      145         150         155         160
184      Ala Gly His Ala Met Val Thr Val Ala Ser Asp Ile Thr Glu Ala Arg
185      165         170         175
186      Gln Arg Ile Leu Glu Leu Leu Glu Pro Lys Gly Thr Gly Glu Ser Lys
187      180         185         190
188      Gly Ala Gly Glu Ser Lys Gly Val Gly Glu Leu Arg Glu Ser Asn Ser
189      195         200         205
190      Gly Ala Glu Asn Thr Thr Glu Thr Gln Thr Ser Thr Ser Thr Ser Ser
191      210         215         220
192      Leu Arg Ser Asp Pro Lys Leu Trp Leu Ala Leu Gly Thr Val Ala Thr
193      225         230         235         240
194      Gly Leu Ile Gly Leu Ala Ala Thr Gly Ile Val Gln Ala Leu Ala Leu

```

DATE: 03/10/2000
TIME: 15:42:36

195					245					250				255			
196	Thr	Pro	Glu	Pro	Asp	Ser	Pro	Thr	Thr	Thr	Asp	Pro	Asp	Ala	Ala	Ala	
197				260					265					270			
198	Ser	Ala	Thr	Glu	Thr	Ala	Thr	Arg	Asp	Gln	Leu	Thr	Lys	Glu	Ala	Phe	
199			275					280					285				
200	Gln	Asn	Pro	Asp	Asn	Gln	Lys	Val	Asn	Ile	Asp	Glu	Leu	Gly	Asn	Ala	
201		290					295					300					
202	Ile	Pro	Ser	Gly	Val	Leu	Lys	Asp	Asp	Val	Val	Ala	Asn	Ile	Glu	Glu	
203	305					310					315					320	
204	Gln	Ala	Lys	Ala	Ala	Gly	Glu	Glu	Ala	Lys	Gln	Gln	Ala	Ile	Glu	Asn	
205				325						330					335		
206	Asn	Ala	Gln	Ala	Gln	Lys	Lys	Tyr	Asp	Glu	Gln	Gln	Ala	Lys	Arg	Gln	
207			340					345					350				
208	Glu	Glu	Leu	Lys	Val	Ser	Ser	Gly	Ala	Gly	Tyr	Gly	Leu	Ser	Gly	Ala	
209		355					360					365					
210	Leu	Ile	Leu	Gly	Gly	Gly	Ile	Gly	Val	Ala	Val	Thr	Ala	Ala	Leu	His	
211		370					375					380					
212	Arg	Lys	Asn	Gln	Pro	Val	Glu	Gln	Thr	Thr	Thr	Thr	Thr	Thr	Thr	Thr	
213	385					390					395					400	
214	Thr	Thr	Thr	Ser	Ala	Arg	Thr	Val	Glu	Asn	Lys	Pro	Ala	Asn	Asn	Thr	
215				405						410					415		
216	Pro	Ala	Gln	Gly	Asn	Val	Asp	Thr	Pro	Gly	Ser	Glu	Asp	Thr	Met	Glu	
217			420					425				430					
218	Ser	Arg	Arg	Ser	Ser	Met	Ala	Ser	Thr	Ser	Ser	Thr	Phe	Phe	Asp	Thr	
219		435						440				445					
220	Ser	Ser	Ile	Gly	Gly	Pro	Cys	Arg	Ile	Arg	Met	Leu	Met	Leu	Lys	His	
221		450					455					460					
222	Arg	Cys	Met	Ile	Arg	Arg	Cys	Arg	Leu	Leu	Ile	Leu	Ile	Arg	Leu	Phe	
223	465					470					475					480	
224	Arg	Ile	Trp	Gly	Ile	Gln	Ile	Ser	Val	Val	Tyr	Ser	Thr	Ile	Gln	His	
225				485						490					495		
226	Pro	Pro	Arg	Asp	Thr	Thr	Asp	Asn	Gly	Ala	Arg	Leu	Leu	Gly	Asn	Pro	
227			500					505				510					
228	Ser	Ala	Gly	Ile	Gln	Ser	Thr	Tyr	Ala	Arg	Leu	Ala	Leu	Ser	Gly	Gly	
229		515						520				525					
230	Leu	Arg	His	Asp	Met	Gly	Gly	Leu	Thr	Gly	Gly	Ser	Asn	Ser	Ala	Val	
231		530					535					540					
232	Asn	Thr	Ser	Asn	Asn	Pro	Pro	Ala	Pro	Gly	Ser	His	Arg	Phe	Val		
233	545					550					555						

234 <210> SEO ID NO 5

235 <211> LENGTH: 1460

236 <212> TYPE: DNA

237 <213> ORGANISM: Escherichia coli

238 <400> SEQUENCE: 5

239	aattctgttg	ctgatgctgc	tgattctcgt	gccagtgata	ttcccggact	tcctacaaat	60
240	ccactgcgct	ttgctgcgtc	cgaggatatct	ttgcatggtg	cgcttgaagt	tcttcatgat	120
241	aaaggggggc	ttgatactct	taactctgct	attggatctt	cgttattccg	tgttgaaact	180
242	cgggatgatg	gcagccatgt	tgctatcggg	caaaaaaatg	gcctcgagac	cactgttggt	240
243	ttaagtgagc	aagagttttc	tagcttacag	tcccttgatc	ctgaaggtaa	aaacaaattt	300
244	gtatttactg	gaggccgcgg	tggcccaggg	catgctatgg	tcacggttgc	ttcagatatc	360

PAGE: 6

VERIFICATION SUMMARY
PATENT APPLICATION US/09/189,415

DATE: 03/10/2000
TIME: 15:42:36

Input Set: I189415.RAW

Line ? Error/Warning

Original Text

93 W "N" or "Xaa" used: Feature required

Gly Asn Ala Ile Pro Ser Gly Glu Leu Xaa A